DOCUMENT RESUME

ED 211 572 TM 810 919

AUTHOR Brittain, Clay V.; Brittain, Mary M.
TITLE Audiovisual vs Paper-and-Pencil Testing of Occupational Competence.
PUB DATE 7 Mar 80
NOTE 13p.; Paper presented at the Annual Meeting of the Eastern Educational Research Association (Norfolk, VA, March 7, 1980).

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS Armed Forces; Attitudes; *Audiovisual Aids; *Comparative Analysis; *Enlisted Personnel; *Job Skills; *Multiple Choice Tests; *Occupational Tests; Reading Ability
IDENTIFIERS Adult Basic Learning Examination; *Paper and Pencil Tests

ABSTRACT

An exploratory application of audio-visual media to the occupation proficiency testing of enlisted soldiers is described. A five-response console which had been developed as a laboratory model was used. A DataQuest machine was adapted for delivery of a multiple choice test of job knowledge. In being tested, the soldier sat in front of the console and listened to the test questions through earphones while viewing supporting visuals projected on a screen. As the DataQuest was connected with a printer, a report of diagnostic/remediation statements was available to the soldier upon completion of the test. Two major conclusions from the study were (1) scores on a multiple choice test of job knowledge were not increased by having the test presented through an audio-visual delivery system rather than in a printed mode; and (2) soldiers expressed strong preference for the machine mediated audio-visual test delivery. (Author/GK)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED211572

AUDIOVISUAL VS PAPER-AND-PENCIL TESTING OF OCCUPATIONAL COMPETENCE.

CLAY V. BRITTAIN
US Army Training Support Center
Fort Eustis, Virginia

MARY M. BRITTAIN
Virginia Commonwealth University
Richmond, Virginia

EASTERN EDUCATIONAL RESEARCH ASSOCIATION
NORFOLK, VIRGINIA
MARCH 7, 1980

TM 810919

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☑ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY
C. Brittain

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

INTRODUCTION and BACKGROUND

This paper describes an exploratory application of audio-visual media to the occupation proficiency testing of enlisted soldiers. It is necessary to begin with a brief description of occupation proficiency testing in the Army.

Since the latenineteen-fifties, MOS proficiency in the Army enlisted force has been assessed through a program of formal testing. The tests which were used until the mid-1970's were known officially as MOS Evaluation Tests, although soldiers referred to them as "pro-pay" tests because they were used as a basis for awarding proficiency (or incentive) pay to soldiers. These MOS tests were, for the most part, multiple-choice tests of job knowledge. The tests came to serve a number of different purposes having to do with personnel management. For example, test scores were used as one factor in selecting soldiers for promotion to the next higher grade, for retention in the Army, for certain educational programs, and - as noted above - for special incentive pay.

In 1974, the decision was made to effect basic changes in Army MOS testing. Tests of job knowledge which had been shaped largely by a norm-referenced logic, were to give way to criterion-referenced tests which emphasized hands-on performance testing. New test nomenclature was adopted. Skill Qualification Testing (SQTs) replaced MOS Evaluation Tests.

An overriding consideration in the change was the goal of having MOS testing be more relevant for the training of soldiers. It was mainly this goal which caused the shift in emphasis toward performance testing. Many of the occupational tasks which soldiers must be trained to perform can be adequately tested only in a performance mode.

In the SQT program, multiple-choice testing has been scaled down, but not eliminated. SQTs still include a multiple-choice test. Thus, we have not gotten away from paper and pencil tests and the use of textual testing materials.

There are at least two major concerns relative to paper and pencil testing. One is that a multiple-choice test in the printed mode may require reading skills in excess of those required by the job itself. Hence, soldiers who are deficient in reading ability may not be able to achieve test scores which are commensurate with their ability to perform on the job. This concern is heightened by the fact that large numbers of soldiers coming into the Army are deficient in reading skills.

In the SQT program, the use of paper-and-pencil testing poses another problem which is a special concern from the point of view of training. This problem has to do with feedback of diagnostic information. There generally is a minimum period of 30 days; and frequently longer, before the results of testing are fed back to soldiers and to the supervisors and commanders responsible for their training.

This was the background against which the present study was undertaken. We were looking for a means of reducing the reading requirements associated with multiple-choice testing and with providing more prompt feedback on test performance.

Experimental SQT. The pilot project was done under contract with the Army by DGA, Incorporated of Silver Springs, Maryland. The project principals were Mr. Henry W. Diambra and Ms. Nancy Gulliford of DGA. The field test was the responsibility of the Army Training Support Center.

The study utilized a machine known as DataQuest which had been developed by the Westinghouse corporation for use in medical history taking. DataQuest presentsquestions in audio and the subject responded by pressing one of three buttons on a console in front of him. There was no visual presentation. The subject could press a repeat button to have a question read again. The repeat capability applied only to the immediately preceding question.

In the present study, a five-response console which had been developed as a laboratory model was used. Other adaptations were made so as to make the DataQuest machine more suitable for delivery of a multiple-choice test of job knowledge. It was modified so as to drive a carousel slide projector, thus making it possible to have supplementary visual materials presented in synchronization with the test questions presented in audio. Thus, in being tested, the soldier sat in front of the console and listened to the test questions through a pair of earphones while viewing supporting visuals projected on a screen in front of him.

DataQuest was also connected with a printer, thus affording the capability of printing a test report on each soldier as he completed the test. The report contained what we called diagnostic/remediation statements. This report could be given to the soldier immediately upon completion of the test.

For purposes of the study a part of an infantry SQT was converted from the printed mode to the audio-visual mode required for delivery by DataQuest. In this conversion, the announcer's speech was compressed so that the time required for the question was reduced to 75 percent of the normal speaking rate. The test contained 48 multiple-choice items covering 10 tasks.

The Field Trial. The DataQuest-mediated SQT was field tested at Fort Bragg, North Carolina, utilizing a sample of soldiers from the 82d Airborne Division. Our goals in the field test were three-fold:

a. To determine whether soldiers, especially those deficient in reading ability, benefited from testing in an audio-visual mode.

b. To assess the reactions of soldiers to this form of testing vis-a-vis a traditional approach.

c. To assess problems associated with this mode of testing.

The present paper addresses mainly the first two objectives. About one month prior to the field test of the DataQuest SQT about 300 soldiers were tested with the Adult Basic Learning Examination (ABLE), Harcourt Brace Jovanich, 1967. The ABLE is described as "a battery of tests designed to measure the level of educational achievement among adults." The battery includes four tests: vocabulary, reading, spelling, and arithmetic. The vocabulary and reading tests from Level II were administered. Level II is targeted to adults achieving at levels comparable to Grades 5 through 8. Only the reading scores are used in the present analysis.

About one-half of the soldiers tested with ABLE also took a printed version of the infantry SQT which had been converted to an audio-visual mode for delivery by DataQuest.

We were interested in comparing performance on the DataQuest version of the SQT with performance on the printed version, and in examining any differential performance in relation to reading ability. As indicated above, our chief interest was in soldiers who were deficient in reading ability. Did such soldiers in fact benefit from taking the SQT in an audio-visual form? Hence our aim was to select a field trial sample so as to obtain disproportionately large numbers of poor readers and we wanted to test each soldier in the field trial with both the printed versionand the DataQuest version of the SQT.

The aims of the field test were not fully realized. In the first place, we did not find the numbers of deficient readers which we had anticipated. The ABLE showed a rather marked ceiling effect for the soldiers tested. In the second place, we were able to exercise only limited control over the testing in the field test sample. This was especially true of the soldiers who were to have been tested with the written version of the SQT following the DataQuest version.

THE RESULTS

Test Scores. Administration of the DataQuest SQT required about 30 minutes per soldier. The field test took two weeks and a total of 101 soldiers were tested with the DataQuest SQT. Of these soldiers, 62 had taken the written version of the test about one month earlier. Our aim was to test the other 39 soldiers with the printed SQT about one month following DataQuest. However, the operational requirements of the units to which the soldiers belonged made it difficult to test the soldiers, and only 19 of them were actually tested. Thus, there was a total of 81 soldiers for whom we had scores on both the DataQuest and the printed versions of the SQT. About three-fourths of these soldiers had taken the printed version first.

Based upon ABLE scores, these 81 soldiers were stratified into three reading groups as follows: below RGL 7.0, 14 soldiers; RGL 7.0 - 8.9, 24 soldiers; RGL 9.0 and above, 43 soldiers. Mean SQT scores for these three sub-groups and for the total of 81 soldiers tested on both versions of the SQT are shown in Table 1. As shown in this table, there is a slight upward trend in SQT scores across the reading levels, with soldiers below RGL 7.0 scoring the lowest and those above RGL 9.0 the highest. In all three reading sub-groups and for the total group, mean SQT scores were slightly higher on the DataQuest version of the SQT.

It may be of interest here to examine the SQT results for the small group of soldiers scoring below RGL 7.0. Of the 14 soldiers in this sub-group, seven took the DataQuest SQT first, followed by the printed version; and the other seven were tested in the reverse sequence. Scores are shown in Table 2. In both sub-groups, scores were slightly higher on the version of the test which was taken second, whether it was DataQuest or the printed version. But the major point is that even in this group of "low readers," scores were not increased by having the test delivered in an audio-visual rather than printed mode.

Attitudinal Data. In addition to the effects of test delivery mode on test performance, we were also interested in assessing the reactions of soldiers to the DataQuest mode of testing.

Two types of attitudinal data were collected. Opinion/attitude questions were included on the last portion of the SQT tape so as to be presented to the examinee following completion of the SQT. The examinee responded to each attitudinal questions by pressing of the buttons on the console, just as he had responded to the test questions.

We also interviewed each soldier immediately after he had completed the DataQuest SQT. The interview was informal and rather loosely structured. We asked the interviewee's permission to tape record the interview. In most instances permission was granted and the interview was recorded.

We will deal first with the structured attitudinal inquiries. Three of these questions addressed points which we thought might present difficulties in the DataQuest mode of test delivery.

The first of these questions had to do with test taking strategies. We were concerned that the DataQuest mode of test delivery did not permit the examinee to skip difficult questions and come back to them later, which presumably is a common strategy in taking multiple-choice tests. One of the questions asked in the post-test segment of the tape was: "When you take written tests, do you skip some questions and come back to them later?"

Overwhelmingly, the examinees answered affirmatively. Ninety-three percent indicated use of this test-taking strategy, 27 percent using it always and 66 percent using it sometimes.

Another concern was that examinees might have difficulty remembering which answer button matched which test alternative. For some test questions, alternatives were visible on a screen until the examinee had indicated his answer. But for other questions, this was not true. It is possible that the examinee would remember the correct alternative but become confused as to whether it was alternative A, B, C, D, or E. We asked the examinees: "Did you have trouble keeping track of which answer choice went with which button?"

About 30 percent indicated that this was a problem, but only 2-3 percent indicated a lot of trouble here.

Still another of our concerns was that examinees would feel hurried by the DataQuest machine. The machine presented the question, then the answer panel lights came on and nothing happened until the examinee responded. Response times were generally very short (one to two seconds). We asked the examinees: "Did you feel that this machine hurried you in taking the test?"

Seventy percent said that they never felt hurried, and 22 percent indicated that to some extent they did feel hurried by the machine. The other eight percent, opting not to answer the question, pressed the "don't know" button.

In order to probe attitudes toward the DataQuest mode of test delivery, the examinees were asked the following questions:

"Did you think this test was fair?"

"Would you prefer to take some other Army tests this way?"

With respect to the question of test fairness, the examinees overwhelmingly responded in the affirmative. Four percent said that the test was unfair and another five percent opted not to answer. The remaining 91 percent felt that the test was fair.

With respect to taking other Army tests in the DataQuest mode, 12 percent either responded negatively or did not respond. The remaining 88 percent indicated a preference for taking other Army tests in the DataQuest mode of delivery.

These favorable reactions toward the DataQuest mode of testing were evident in the interview data. The interviews, which were conducted quite informally by the senior author and another research psychologist, were tape recorded and subsequently transcribed. Two other persons read the interview protocols and jointly categorized them as favorable, neutral, or negative toward the DataQuest mode of testing. The protocols were classified as follows:

Favorable	84 percent
Neutral	11 percent
Negative	6 percent

One has only to read through the protocols to be convinced that this is a substantially accurate summary of what was said in the interviews. The excerpts given below seem somewhat representative. A theme that runs through the interview protocols is that not having to read the test makes it easier for the examinee.

I think it's more fair, more understandable. It's about the only test I've understood that much.

I'd like to take tests like that all the time. It's a lot easier to understand.

I think it was easier that way. You understand it better. For myself, it was.

It's easier for a person, you know, to comprehend it more better when he's listening than when he's reading. . . it gives him more time to think about it than if he reads it, and it keeps a person from falling asleep. . . This video tape . . . keeps you up. Well, I did. I got more interested. I think it's better.

I'm not that good in reading English, so I like this test very good.

It was all right! That was pretty nice . . . Pretty slick way of taking a test . . . It was easier to answer the questions when the answers were on the screen, because my memory's not too good at things like that . . . It was a good test.

That's one nice machine. I say the machine was fantastic. It's something else. The people I know personally, if they took a test like this, they'd love it. They'd actually love it. They definitely would.

My score would be much higher. 'Cause, like I say, reading's fantastic, but that machine got something over that reading. It's right there, instead of trying to think about what you read, and try to comprehend, and all that.

I say it's fantastic. I never seen a machine like that. I like that. What you're already learned in training . . . will come right back to you just like that. . .. Some part of the question he asked me, if I had read it, I definitely wouldn't have got it.

I think it's an improvement. It's better than to just give somebody about so many pages of paper and then let them sit there and get bored taking the test.

Lots of guys out in the field--people that don't understand what they're reading too well--would be excellent for them.

Just takes a lot of presssure off you, too. You're more relaxed, and that's the kind of atmosphere that's really needed when you're taking a test. Well, a lot of pressure on you, people looking, breathing down your neck. A great improvement.

Listening to the story is a lot better than reading it because if you listen, you generally pay more attention to it. Where, if you read, your mind might tend to drift away from what you're really reading about, you know.

I like it better; I actually do. Sometimes it get boring reading. Like some of those questions on some of them other tests, you got to fill in a little square. I used to do it--and all my friends--we used to get fed up with it and just go down and mark something.

I like this a lot better. It's less boring. I just feel like I can answer it better. I can understand it a little bit more.

It's a whole lot better than just sitting there, be given instructions, and you have to write this, color this dot, color that dot, pass the papers up, and stuff like this; when all you got to do is sit there and push buttons on this one--and it's a whole lot easier.

I can hardly read something just to read it out. I don't understand it too good, if it is like that. If I had someone telling me questions, I understand it better. It would be a little fun doing it that way, than just reading it like that.

CONCLUSIONS AND DISCUSSION

The two major conclusions from the study are:

1) Scores on a multiple-choice test of job knowledge were not increased by having the test presented through an audio visual delivery system rather in a printed mode.

2) Soldiers expressed strong preference for the machine mediated audio visual test delivery.

The strong positive reactions of soldiers to the DataQuest mode of testing is interesting and perhaps even surprising. The DataQuest did not allow the use of what apparently is a commonly used strategy in taking multiple-choice tests; namely, skipping difficult questions and going back to them later. Also a number of soldiers reported that to some extent they felt hurried by the DataQuest machine. Soldiers also reported experiencing some difficulty in keeping track of which answer choices went with which response buttons. Perhaps the opportunity to have each question repeated as many times as necessary neutralized these potential problems in the machine mediated test. Also the rather marked preference for the DataQuest SQT over the printed version was not reflected in test scores. Although they preferred taking the SQT in the DataQuest mode, the soldiers scored just as well on the printed version.

TABLE 1

Scores on Data Quest and Printed SQT in relation to Reading Ability

RGL	N	DATA QUEST		PRINTED	
		MEAN	SD	MEAN	SD
Below 7.0	14	30.5	3.96	29.7	5.28
7.0-8.9	24	33.5	4.90	32.6	3.80
9.0 & above	43	34.8	4.72	34.3	4.87
Total	81	33.7	4.86	33.0	4.96

TABLE 2

Test Scores in Relation to Test Sequence for Lowest Reading Group.

SEQUENCE	N	DATA QUEST		PRINTED	
		MEAN	SD	MEAN	SD
Printed SQT First	7	30.0	5.1	26.9	5.8
Data Quest First	7	31.0	2.7	32.6	2.7